

# ***A User's Guide to Semio Taxonomy and SemioMap***

*Version 4.1*

*February 2001*

The information in this document is subject to change without notice and should not be construed as a commitment by Semio Corporation. Semio Corporation assumes no responsibility for any consequences resulting from errors that may appear in this document.

For further information regarding this document, contact:

Semio Corporation  
1730 S. Amphlett Blvd.  
San Mateo, CA 94402

Phone: (650) 638-3330  
Fax: (650) 638-3339  
[www.semio.com](http://www.semio.com)  
[Info@semio.com](mailto:Info@semio.com)

#### GOVERNMENT USE

All Semio software products are commercial items as defined at 48 CFR 2.101 (OCT 1995), consisting of commercial computer software and commercial computer software documentation. Consistent with 48 CFR 12.212 and 227.7202-1 through -4 (JUNE 1995), any license or sublicense to United States Government end users will be made with only those rights as are granted to all other end users pursuant to Semio Corporation's standard end user license terms.

Copyright © 2000, 2001 Semio Corporation

All Rights Reserved.

SemioMap® is registered in the U.S. Patent and Trademark Office. Semio Taxonomy and Semio Tagger are trademarks of Semio Corporation. Netscape is a registered trademark of Netscape Corporation. Microsoft, Windows95 and Windows NT are registered trademarks of Microsoft Corporation. All other product and trade names mentioned are registered trademarks or trademarks of their respective companies.

---

# Contents

Preface	v
About This Document	vi
<b>Chapter 1</b> <i>Introduction to Taxonomies and Information Discovery</i>	<b>1</b>
Searching versus Discovery	2
Structures That Support Discovery	4
Concepts instead of Documents	5
<b>Chapter 2</b> <i>Introduction to Semio Taxonomy</i>	<b>9</b>
The Semio Taxonomy Solutions	10
Customizable Taxonomies	10
Multiple Taxonomies for Multiple Audiences	11
Unique Features in Semio Taxonomies	12
Ways of Using Semio Taxonomy	14
<b>Chapter 3</b> <i>Navigating Taxonomies</i>	<b>15</b>
Browsing the Taxonomy	16
First Page	16
Subcategories	17
Concepts	17
Related Terms	18
Documents	18
Searching for Specific Concepts or Words	21
<b>Chapter 4</b> <i>Navigating Concept Maps</i>	<b>23</b>
Understanding Concept Maps	24
Starting SemioMap Discovery	27
Navigating Concept Maps	30
Understanding Map Shapes	34
Viewing Source Documents	37
Index	39



---

# *Preface*

---

Semio Corporation's software creates browseable structures of the concepts in text-based content, making it easier to find and discover relevant information online. SemioMap creates navigable, graphical maps of text-based content, and Semio Taxonomy creates custom topic hierarchies for corporate portals and web sites.

This document is a guide to using Semio Taxonomy and SemioMap. It is intended for Semio Taxonomy users and SemioMap users, which may include

- General users of the taxonomies on intranet or Internet sites
- Site owners
- Knowledge workers

This document assumes that the Semio solution is installed and the taxonomies or maps been built.

Information on building and tuning taxonomy and map structures may be found in other documents, available from Semio Corporation.

---

## *About This Document*

This document is organized as follows:

<b>Chapter</b>	<b>Title</b>	<b>Contents</b>
1	Introduction to Taxonomies and Information Discovery	This chapter provides a high level overview of taxonomies and information discovery.
2	Introduction to Semio Taxonomy	This chapter provides an overview of how Semio Taxonomy works.
3	Navigating Taxonomies	This chapter describes the details of navigating the Semio taxonomies.
4	Navigating Concept Maps	This chapter describes how to navigate concept maps in SemioMap

---

# *Chapter 1 Introduction to Taxonomies and Information Discovery*

---

At no time in history have individuals had access to so much information, virtually on-demand. With the growth and expansion of the Internet and online information sources, a network connection can practically deliver the world to the desktop.

But more information, alas, does not necessarily lead to more knowledge – knowledge of the kind that makes our lives easier or our businesses more efficient and profitable. The large quantities of data available can actually impede our search for knowledge:

- Users confronted with enormous results from keyword searches can find it difficult to filter out unimportant or irrelevant information.
- The large volumes of information can obscure the relevant relationships between pieces of information. In the world of structured information, data mining technologies help display these relationships. For unstructured, text-based information, there are no such tools.
- In rapidly-changing industries such as technology and science, knowledge experts have an enormous task in remaining up-to-date.

*Structured data*, represented by transaction-oriented data stored in corporate databases, is increasingly accessible through new classes of decision support tools,

including data mining tools and analytic applications. But much of the information flooding us everyday is *unstructured data* – text that we must read and interpret.

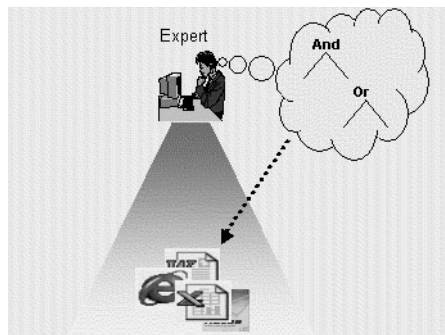
Semio software is built to address this problem. Semio Taxonomy supplements traditional keyword searching technology by providing browseable *structures* built from the concepts in the source information. By providing browseable structures in the form of maps or topic hierarchies, Semio's software enables information *discovery*, giving users better access to large amounts of text-based information and helping them explore this information and uncover new relationships between concepts in the text.

The following sections describe these ideas in more detail, in the context of traditional keyword search engines and familiar structures for web-based information.

### Searching versus Discovery

*Searching* as we know it today is the process of entering a keyword and searching against indexed information. Searching provides a necessary point of access into unstructured data.

Keyword searching is essentially a *top-down* process, driven by a user's specific formulation of a search.



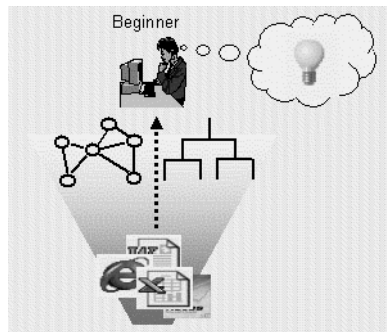
---

A keyword search can be the most efficient way to find information when you are very clear about what you need to find and can create a precise query. However, keyword searching has a number of constraints, which become significant as the amount of information being searched grows.

- Assembling a good search term can be difficult.
- Making sense of a large number of results is challenging.

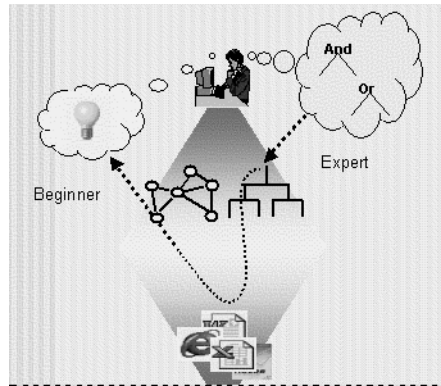
Keyword searching returns a *listing of results*; it does not return a structure that users can explore and learn from. This is the realm of information *discovery*.

Discovery is essentially a bottom-up process; the unstructured data is processed to expose key concepts and relationships. Users do not have to know what data is available, or to have a firm idea of what they are looking for. In the discovery model, the user may be a beginner in the subject area.



Discovery provides structures by which the user can explore and discover information. This is significantly different than simply presenting a list of search results. Instead, the user actively examines and interacts with the data, gaining an overall understanding of what is available, discovering connections between subjects, and potentially finding relevant information that the user otherwise did not know was there.

Ideally, searching and discovery should both be present in the same system, as an expert in one area may shift his focus slightly and become a beginner exploring a related topic.



Implementing both keyword search and discovery creates a rich learning environment, in which it is easier for users to find the valuable information they need from a wide variety of potential sources.

## Structures That Support Discovery

While keyword searching delivers a flat listing of results (sometimes ranked by apparent relevance), individuals *discover* information by browsing *structures* that represent the data in some way. A structured view into text-based sources supports information discovery by:

- presenting a high-level view of information
- offering paths for exploring information
- helping the user discover new associations and relationships in the data

Examples of structures include:

- A *bunch* of information (or an unstructured structure). Examples include the home pages of many large vertical portals. The page is filled with a number of different topic area links – current headlines, sports news, local weather, links to financial information, etc. This gives you a broad sense of what information is available on the site.

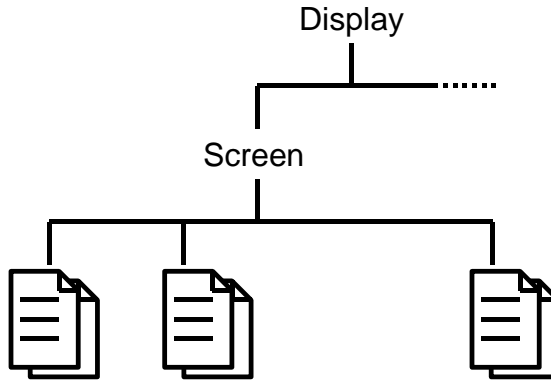
- 
- An *index* or *site directory*. These provide a listing of what is available on the site, typically in alphabetical order. One analogy is a book index.
  - A *directory*. A directory is a fairly flat classification that allows a user to look at lists of related objects. Examples include alphabetical listing of names and addresses, or annual reports by industry sector. Directories may be combined with keyword search to help people target searches.
  - A *taxonomy* is a systematic classification of a conceptual space. For example, in biology a taxonomy of living things separates mammals from birds and spiders from insects, based on structural differences.

Semio Taxonomy combines a high-level topic hierarchy built specifically for the site with a conceptual taxonomy created automatically from the content itself. Semio's taxonomy uses phrases (identifying key concepts) within the text as the elements of classification. These two entities are linked together using knowledge classification rules and Semio's knowledge engineering expertise. Once created, the entire structure can be fine-tuned for optimal results, and automatically refreshed and recreated as the source content changes.

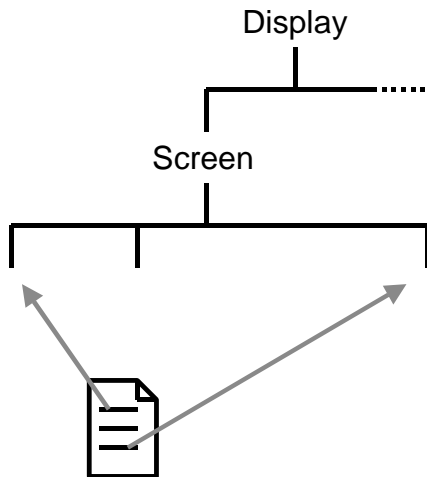
## Concepts instead of Documents

One important distinction between Semio Taxonomy and simple categorization tools is that Semio Taxonomy works with the *concepts* within documents.

Many categorization schemes try to put each document into one or more distinct categories.



Semio Taxonomy, however, exposes the *concepts* within the documents; a single document may be accessed from many different parts of the taxonomy if its content matter is relatively broad.



By categorizing the concepts in the documents instead of the documents themselves, Semio Taxonomy provides insight into diverse document sources, as well as an overview of the concepts covered in the collection.





---

## *Chapter 2 Introduction to Semio Taxonomy*

---

This chapter provides a high level overview of how Semio Taxonomy builds taxonomies from text-based information.

## *The Semio Taxonomy Solutions*

Semio software uses patented lexical extraction and clustering techniques to create navigable structures of unstructured information, enabling information discovery. The process by which this happens is as follows:

- Semio's core software (Semio Tagger) collects information from a variety of sources, which may include Lotus Notes databases, web sites, XML files, local files, and most other common data sources.
- A patented lexical extractor extracts key concepts from these text sources.
- A database builder builds an index of the occurrences and co-occurrences of key concepts in the text.
- The taxonomy builder builds browseable, searchable directories from these databases, linked to customized high-level directories created for each installation.

The resulting topic hierarchy (taxonomy) is suitable for a wide range of users; most users are already familiar with the model of browsing Internet directories. Semio taxonomies are ideal for improving access to information on corporate portals and other information-provider sites.

### **Customizable Taxonomies**

Each Semio taxonomy is created from two different layers:

- The top level *ontology* provides the high level categories for the specific site. For example, an internal corporate portal may have categories for different divisions and for personnel matters, while a site hosting primarily medical information will have a very different set of high level categories.
- The collection of key concepts is created from the text itself.

The Semio software automatically links these key concepts to the top-level ontology using parameters configured during the taxonomy creation. You can update the taxonomy with new content at any time, and Semio provides tools for tuning and adjusting the taxonomy.

This architecture combines the best of automation and human expertise:

- Human expertise determines the high-level categories that are relevant to the site's users. These may be adjusted over time.
- Automated software creates the taxonomies from the source text, processing in minutes what would take an individual months to read and categorize.
- Human expertise fine-tunes the results for optimal utility.
- Automated software maintains the taxonomy, without manual intervention, document tagging or categorization.

The result is a taxonomy that is tailored to the specific needs of the site and yet is easy to update as sources change.

### **Multiple Taxonomies for Multiple Audiences**

Sometimes a site serves a diverse audience with different needs for the same data. For example, individuals in finance need different kinds of information or views into information than individuals in marketing.

One approach is to create a broad taxonomy that includes all of the concepts anyone might need from a broad document base. Another approach is to create specialized taxonomies for specific audiences, representing different views of the same information.

This second solution is not implemented in a widespread fashion today because many sites use manual methods to categorize documents and put them into a hand-maintained directory. But with Semio's configurable and automated taxonomies, creating multiple views into the data is not only possible, it is easy. You can create multiple taxonomies from the same source data and maintain them automatically.

This provides another mechanism for making it easier for the users to find what they need.

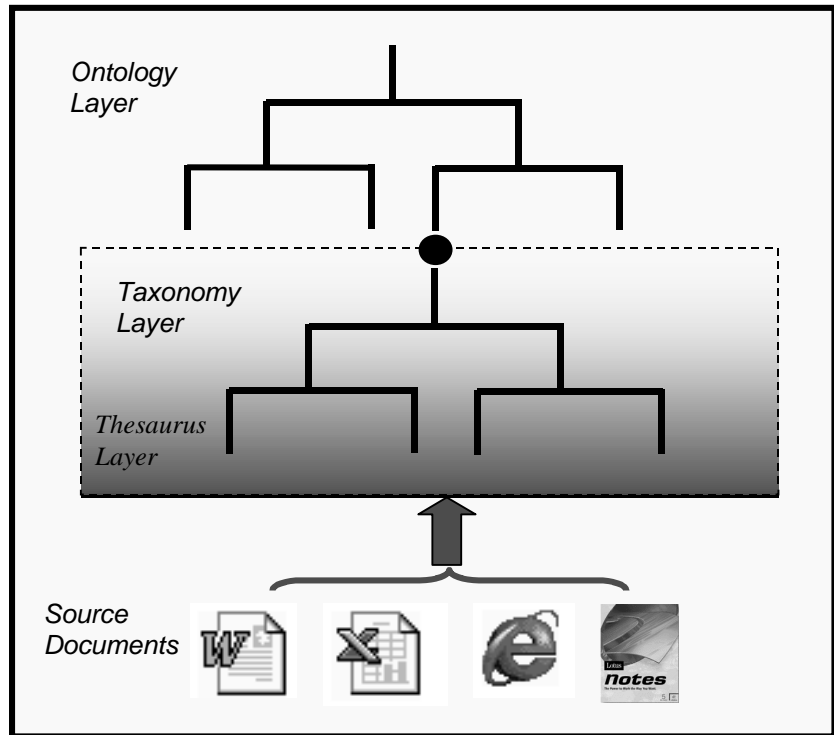
## *Unique Features in Semio Taxonomies*

The overall structure and use of the Semio taxonomy should be familiar from other web-based structures, such as Lycos and Yahoo! directories. However, the Semio taxonomies have several key differences:

- The lower levels of the taxonomy list *concepts* (typically phrases) derived directly from the text. These provide unique insight into what is discussed in the documents.
- The Related Terms sections list concepts that occur together with the selected topic in the source texts. These help identify connections between ideas in the text, much like a thesaurus helps identify related or similar words.
- Semio Taxonomy automatically displays the most relevant documents for a selected concept.

The Semio taxonomy is comprised of two key components: the higher level categories (ontology) created as part of the taxonomy design, and the lower-level taxonomy structure and cross references, automatically generated from the text.

In the graphic below, the ontology layer is the high level categories and subcategories for the taxonomy. This is created manually.



The taxonomy and thesaurus layers are generated by the Semio Taxonomy software and automatically linked together using Semio's lexical tools and knowledge library. Once the structure is completed, it can be tuned and updated quickly and easily. New content can be added automatically, without manually changing the taxonomy structure.

## *Ways of Using Semio Taxonomy*

Most Internet users are already familiar with the concept of browsing through linked directories of information. This is a model created and perpetuated on the general purpose web search engine sites such as Lycos and Yahoo.

Semio's taxonomies are similar in their overall appearance and use to these familiar structures. Users should be able to start using Semio taxonomies to find information immediately, without any specialized training.

At the same time, a deeper understanding of how taxonomies work is helpful for many individuals. For example,

- A knowledge worker using the taxonomies extensively to survey new information in the field could benefit from a solid understanding of Semio's related concepts.
- A site administrator can use a thorough knowledge of the taxonomy structure to understand and evaluate the overall scope of the document sources.
- A content expert could use the SemioMap component to survey information in a new field quickly and efficiently.

The following chapter describes in more detail how to navigate Semio taxonomies and Semio concept maps.

---

## *Chapter 3 Navigating Taxonomies*

---

A Semio taxonomy is a hierarchical structure of the concepts in the available text, with the most general concepts in the top level categories and the most detailed listings further down.

There are several ways to use the Semio taxonomy structure:

- Browse the taxonomy by selecting categories and subcategories and following links.
- Search for a specific concept or keyword in the structure.
- Examine the concepts in context in the source documents and refine your search to deliver the documents you need.

## *Browsing the Taxonomy*

The Semio Taxonomy page has four main components

- Categories
- Concepts - concepts extracted from the text and categorized
- Related Terms - concepts that are somehow related to a selected concept in the text. These are useful for finding connections and cross-referencing.
- Documents

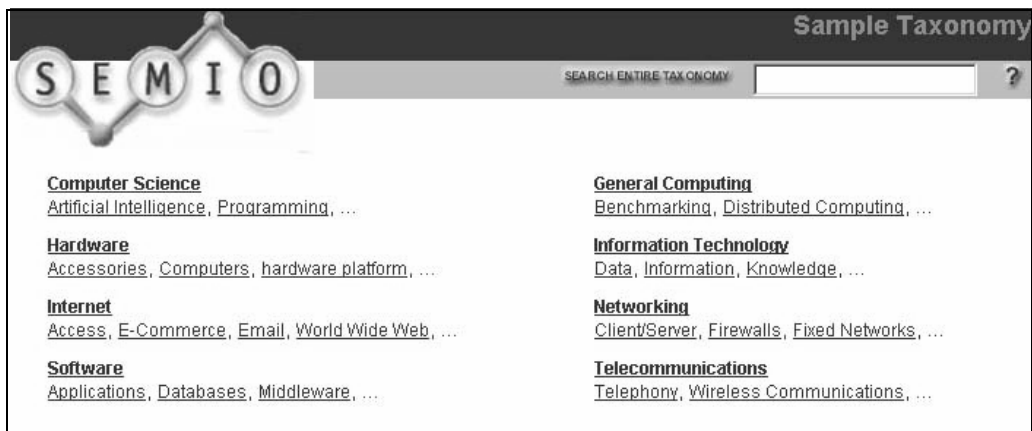
Every taxonomy page will have one or more of these components. Expand/hide buttons in the top right of the different bars let you vary your display.

The top of the screen lists where you are in the taxonomy at any time:

Top:Investing

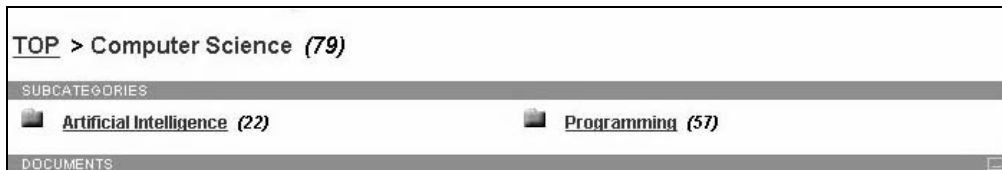
### **First Page**

The first page of the taxonomy contains predefined categories at a high level. You can select a category to browse its contents, or select one of the subcategories displayed on this screen to jump directly to that level.



## Subcategories

The Subcategories section lists any subcategories within the selected category. These categories are created by the taxonomy designer.



Clicking on a category links to a page for that category, which may also have subcategories.

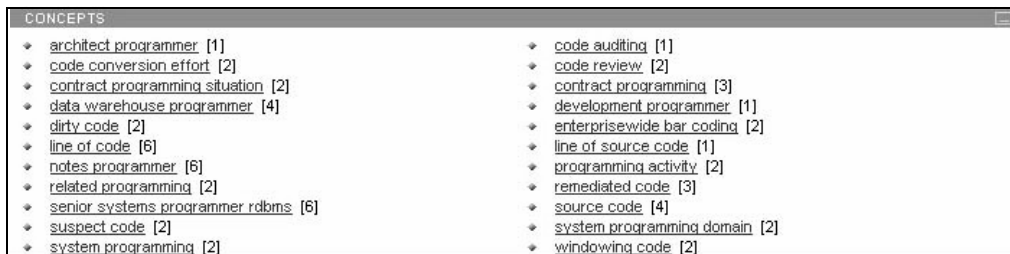
**Numbers.** Parenthetical numbers after the categories indicate how many documents are referenced by each subcategory (and all of its subcategories).

**Other Paths.** If a category exists in more than one location in the taxonomy, the Other Paths icon appears next to it. Clicking on this button displays the other taxonomy location(s) containing this category.

## Concepts

Semio Tagger applies patented methods to determine which phrases are important concepts in the source text. It extracts these concepts and automatically links them into the category structure of the taxonomy.

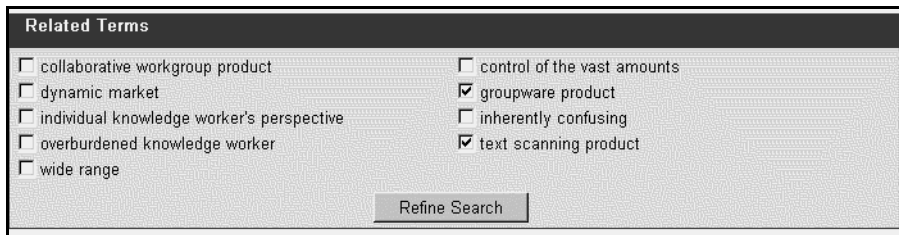
The Concepts section lists the relevant concepts for the selected category. These concepts are listed as they appear in the text.



Clicking on a Concept displays a list of concepts related to the current phrase in the text. A parenthetical number next to a concept indicates how many documents are linked to the concept.

### Related Terms

Once you select a concept, you see a list of the related terms to that concept. This is useful for narrowing down your document listing. Selecting one or more terms from this list and clicking the Refine Search button displays only the documents containing these terms.



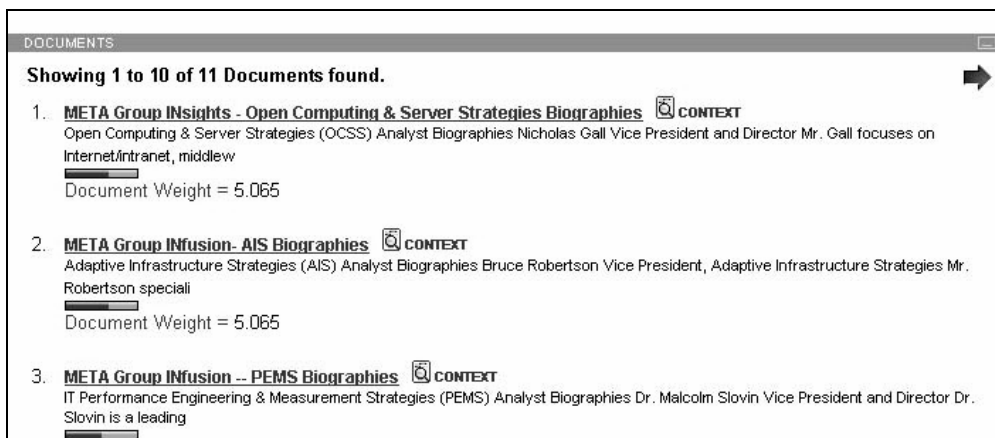
Related Terms	
<input type="checkbox"/> collaborative workgroup product	<input type="checkbox"/> control of the vast amounts
<input type="checkbox"/> dynamic market	<input checked="" type="checkbox"/> groupware product
<input type="checkbox"/> individual knowledge worker's perspective	<input type="checkbox"/> inherently confusing
<input type="checkbox"/> overburdened knowledge worker	<input checked="" type="checkbox"/> text scanning product
<input type="checkbox"/> wide range	

Refine Search

**Selecting a linked concept.** If the related concept is underlined, it means it appears somewhere else in the taxonomy. Clicking on it displays another location.

### Documents

Once you select a concept, Semio Taxonomy displays a list of the documents that contain that concept. The most relevant documents are generally listed first.




The document list includes linked entries to the source documents, along with the first line or so of the text. Selecting a document displays the document (in some cases, launching plug-ins to access the source document, or connecting to a web site).

You can also see related documents at the Subcategory level by expanding the Documents section; simply click the Expansion icon in the top right corner of the Documents title bar. This shows all of the documents linked to the selected subcategory. Clicking on this box again hides the document listing.

**Relevance bar/document weight.** Documents are generally listed with the most relevant citations first. The graphical relevance bar and the document weight number show how important the selected concept is in the document. The number associated with is from 0-10, with 10 being the most important.

**Viewing concepts in context.** Semio Taxonomy also provides a way for you to get a quick idea of how the concept is treated in the document without reading the entire document. Click on the Context icon to see the paragraphs in the document that contain the selected phrase(s).

Semio Context Display

Selected Phrase(s): computer science

... Dr. Slovin is a leading authority on application development, maintenance, and re-engineering methodologies and processes. In addition, he is a noted expert in the application of metrics to software process and product improvement. Prior to joining META Group in October 1996, he was a principal with **Computer Sciences** Corp.'s (CSC) Consulting and Systems Integration division and principal architect of CSC's Center of Excellence for I/S Asset Management. The Center examined, integrated, and developed new and emerging technologies, techniques, and tools for improving productivity and quality in legacy environments. Dr. Slovin holds a Ph.D. in Zoology (with work in the areas of **Computer Science** and Statistics) from Indiana University and was an adjunct professor at the University of Maryland in **Computer Science**. ...

... Dr. Bohner is a leading authority on software maintenance and re-engineering, and is an internationally recognized expert on software impact analysis. Before joining META Group in January 1997, he was a lead scientist with Mitretek Systems (MITRE Corporation), heading their year 2000 compliance program which served public sector clients. Dr. Bohner was responsible for leading software modernization, process improvement, and business process re-engineering efforts. He holds B.S. and M.S. degrees in **Computer Science** from the University of Maryland and Johns Hopkins

This opens a separate browser window. You can read the context to determine if this document is helpful.

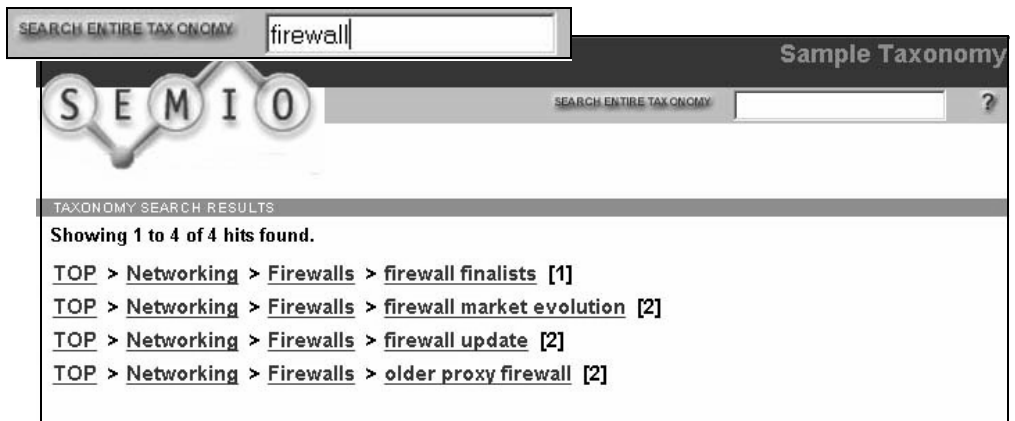
**Refining your search.** If the document list is too long, you can select refine your search using the Related Terms box. Simply select one or more checkboxes in the Related Terms area and click **Refine Search**. This displays only the documents containing the related terms in addition to the selected concept.

## Searching for Specific Concepts or Words

You can search for a specific item within the taxonomy structure by entering a word or phrase in the Search Entire Taxonomy box and pressing Enter.

A rectangular search box with a grey border. On the left, there is a label "SEARCH ENTIRE TAXONOMY" in a smaller, lighter box. To the right of the label is a larger white input field.

If your result appears in more than one location in the taxonomy, you see a screen with the different taxonomy areas.

A screenshot of a web application interface. At the top, there is a search bar with the text "SEARCH ENTIRE TAXONOMY" and a search button. Below the search bar, there is a header area with the text "Sample Taxonomy" and a search bar. The main content area is titled "TAXONOMY SEARCH RESULTS" and shows "Showing 1 to 4 of 4 hits found." Below this, there are four search results listed in a table-like format. Each result shows a path from "TOP" to "Networking" to "Firewalls" to a specific topic, followed by a count in brackets. The results are: "firewall finalists [1]", "firewall market evolution [2]", "firewall update [2]", and "older proxy firewall [2]".

TAXONOMY SEARCH RESULTS	
Showing 1 to 4 of 4 hits found.	
TOP > Networking > Firewalls > firewall finalists	[1]
TOP > Networking > Firewalls > firewall market evolution	[2]
TOP > Networking > Firewalls > firewall update	[2]
TOP > Networking > Firewalls > older proxy firewall	[2]

Clicking on a listing displays that taxonomy page.



---

# *Chapter 4 Navigating Concept Maps*

---

SemioMap lets you view and navigate three-dimensional concept maps of the source text. These maps can provide great insight into a document collection.

Concept maps are particularly useful for a number of tasks:

- Surveying quickly the information contained in a large, text-based collection by viewing the relations of concepts.
- Getting up-to-speed quickly in a new area by exploring its key concepts.
- Discovering unexpected relationships between concepts in the sources.

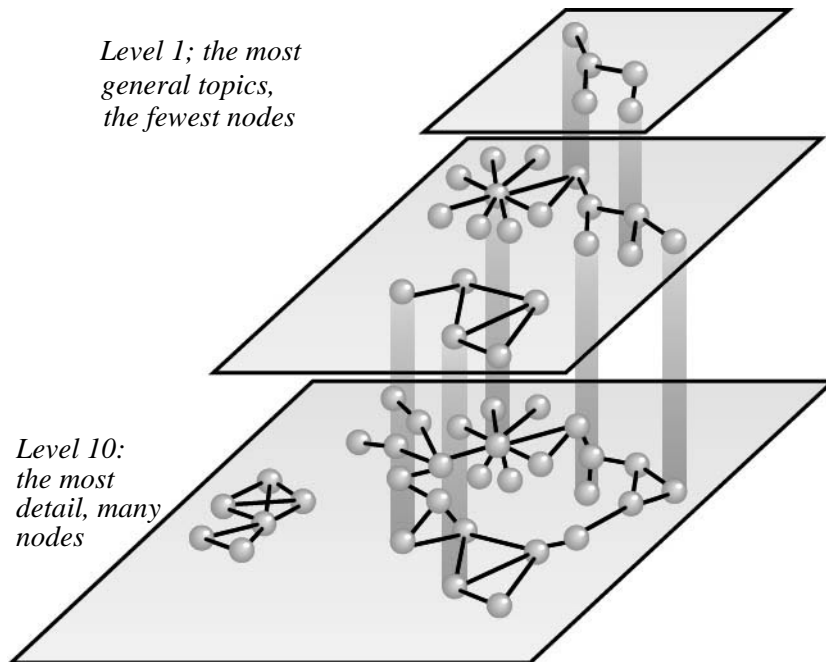
This chapter describes how to use these maps to glean meaning from the Semio sources.

## Understanding Concept Maps

A *concept map* is a three-dimensional representation of the topics in the source collection.

Topics in the source text are represented by *nodes* in the map. A node is a keyword or phrase that represents an important idea from the source document. It comes either from a list provided during database creation or as the result of Semio's linguistic analysis.

Lines between nodes indicate that these topics are related to each other (that they occur near each other in the source document).



---

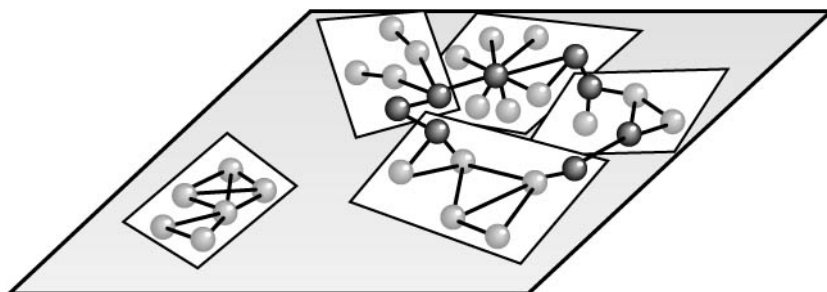
*This diagram illustrates a concept map. Nodes (representing topics) occur at different details levels in the concept map. The nodes that occur most frequently appear at the highest, most general level. Other nodes only appear at lower levels.*

**Levels.** A concept map has up to ten different *levels* of detail. Only the nodes that occur together very frequently appear at the most general level (level 1). This level displays the most prominent topics in the collection, or the “big picture” of the collection.

Lower levels include topics that occur less frequently, and represent a more detailed view of the source text.

Topics that occur at the most general level also occur at the detailed, lower levels.

**Shared nodes.** A node may be linked to several maps; this is called a *shared node*. The illustration below shows how nodes are shared between different maps.



---

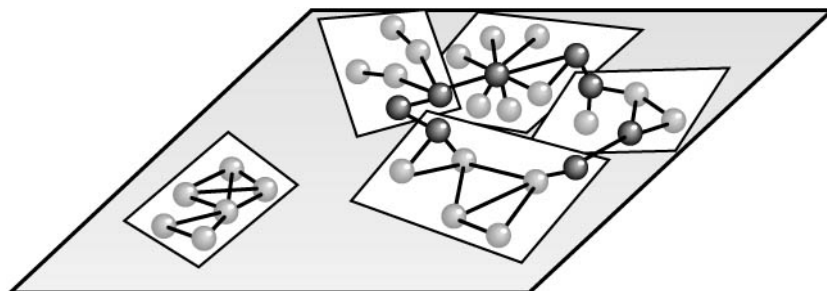
*The highlighted nodes are shared nodes; they are shared between multiple maps.*

Double-clicking on a shared node presents the other map(s) with which the node is linked.

**Maps.** At each level of the concept map are multiple maps. SemioMap displays one map at a time. When you double-click on a shared node, SemioMap displays this node’s links to other maps. You can move directly to a map by double-clicking on its entry in the **Maps** scroll box. Maps are named for their central node.

Traveling through a single level of the concept map in this way is called *panning* through the map.

**Worlds.** The diagram below shows one map not linked to the other maps; this is a separate *world*. (In the diagram, this level has two worlds.) The maps in one world are not connected at this detail level to the maps of another world.



---

*The map in the lower left corner is in a different world than the other maps at this level.*

Move to another world by double-clicking on an entry in the **Worlds** scroll box.

As you navigate the concept map, you will pan through topics on a level, and zoom down and up through the detail levels to find the information you need.

## Starting SemioMap Discovery

There are two ways to start exploring a database.

- Start SemioMap directly (for Windows systems only).
- Start your web browser and enter the URL for SemioMap Discovery.

Contact your Semio administrator for instructions on starting SemioMap in your environment. The general instructions below are typically appropriate.

*Note:* Improve the performance of SemioMap by installing the Semio Java classes on the machines you use to run SemioMap. See your Semio administrator for details.

**Starting SemioMap from Windows.** From a Windows system:

1. Make sure the web server is running on the host system, or start the Semio Admin Server.
2. If you have installed the Semio client software on your system, then start SemioMap from the Windows Start menu:

**Start | Programs | Semio Software | Semio Discovery**

This automatically launches the web browser and displays the SemioMap catalog.

Otherwise, start the web browser and enter the URL for SemioMap. For example,

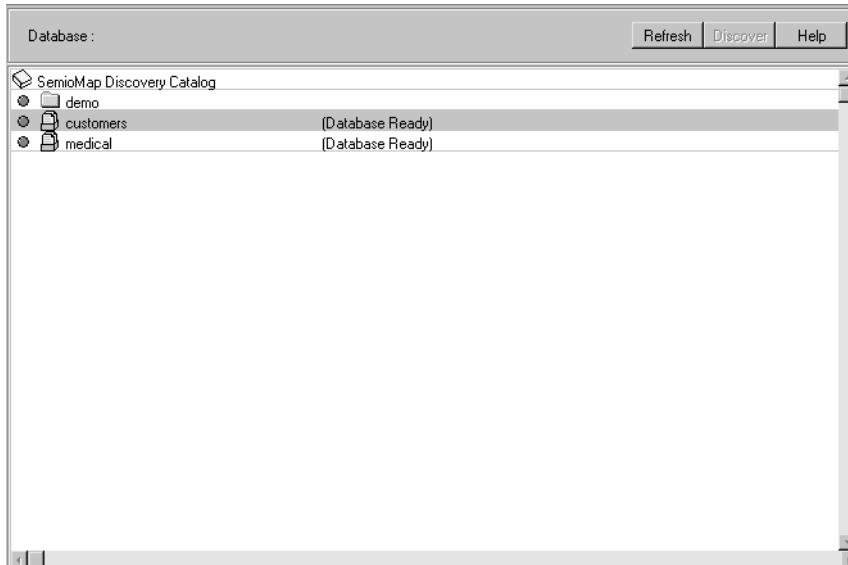
`http://semio/discover.cgi`

**Starting SemioMap from Solaris.** From a Solaris system:

1. Make sure the web server is running on the host system, or start the Semio Admin Server.
2. Start the web browser software.
3. Enter the URL for SemioMap, with the discover.cgi extension. For example:

`http://semio/discover.cgi`

**The Discovery Catalog.** When you start SemioMap, you see the Discovery catalog.

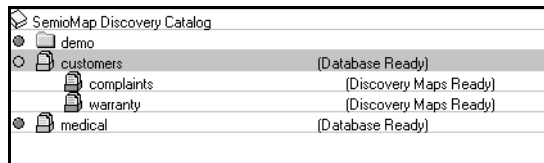


---

*The Discovery catalog. This screen shows one catalog group, called Demo, and two databases. The circle by an object indicate that it has more objects belonging to it. Double-click on a folder or database to see what is below it.*

A catalog group is a group of databases and maps, and appears as a folder. Demo is a catalog group in the example above. To see its contents, click on the small circle next to its name, or double-click on the folder.

A database is ready if it has the “Database Ready” indication next to it. You cannot navigate a database; you must instead navigate a *map* belonging to a database. To see the maps that belong to a database, click on the circle next to its name, or double-click on the database.



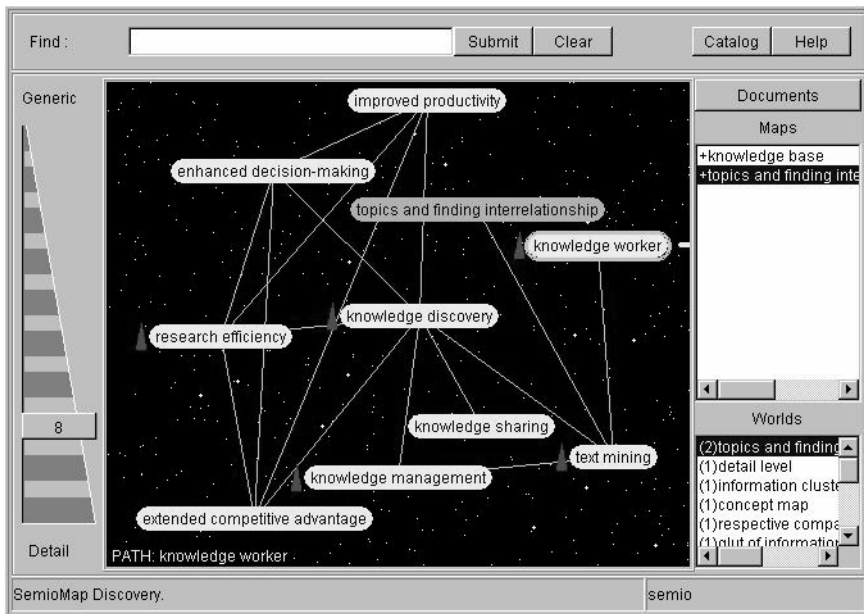
Maps that can be browsed will have “Discovery Maps Ready” next to them.

To start exploring a concept map, select the map and click the **Discover** button, or simply double-click on the map.

## Navigating Concept Maps

SemioMap displays the first map in the concept map for the selected database. You may need to use the scroll bars in your web browser to see all parts of the SemioMap screen.

You can hide navigation bars on the browser to make more of the screen accessible for the concept map.



There are many components to this screen. The map, at center, displays concepts in the text collection. The **Maps** and **Worlds** scroll boxes display other maps available at this level. The **Detail** slide bar at left provides a way of moving through detail levels and shows which level you are currently seeing. The **Find** box moves to a specific topic.

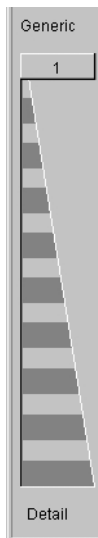
The following sections describe this screen in more detail.

**Nodes and Links.** Nodes with dark borders are *shared nodes*; they are linked to multiple maps at this detail level. You can see small nodes near the shared node; these are referred to as *satellite nodes*, and indicate that the node is shared with another map. Double-clicking on a shared node displays its links to other maps.




---

*“Storage device” is a shared node, with links to other maps.*



**Detail slide bar.** To change levels, select a node and use the Detail slide bar.

- Click once below the pointer to move down one level (to more detailed maps).
- Click once above the pointer to move up one level (to less detailed maps).
- Slide the pointer to move to a specific level.

SemioMap uses the currently selected node to travel through the detail levels. By default, SemioMap selects the node with the most links (the leading node) in the current map. To see a different concept at another level, select a different node before using the detail slide bar.

A red arrow by a node indicates that it occurs at a higher, more general level. If you select the node, you can use the Detail slide bar to see this node at a higher level. (You can travel to a lower, more detailed level with any node.)

**Maps scroll box.** To the right of the map is a scroll box of maps. Each entry represents a map available at this level; these maps are connected, either directly or

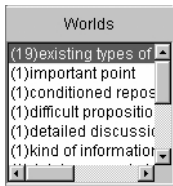
indirectly, with the map that you are seeing. The maps are named for the central node of each map.



To move directly to one of these maps, double-click on its entry in the **Maps** scroll box.

A plus sign by a list entry indicates that this map has at least one node that also exists at a higher level in the concept map; you can travel upward from this map.

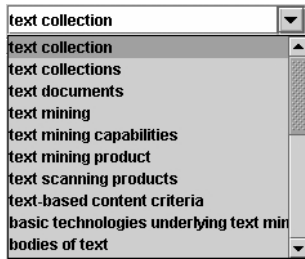
**Worlds scroll box.** At the bottom right of the map display is a scroll box listing other *worlds* at this level. A world is a collection of connected maps. Different worlds are not connected in any way at the current level.



A world is named for the central map in the world. The number by the entry indicates how many maps are in each world.

Nodes are not linked across worlds; to access a different world, double-click on an entry from the **Worlds** list. This displays the central map for the selected world. The entries in the **Maps** scroll box are refreshed to represent only the maps in this selected world.

**Find box.** The **Find** box lets you direct your exploration of the concept map through the use of a specific word or phrase. Enter the word or phrase and click the **Submit** button. If there is one exact match, SemioMap displays the map that contains the node, at the highest level in which the node appears.



If there is no exact match, SemioMap returns all possible matches in the **Find** box. Click on the **Find** box to see a pop-up menu with the possible matches.

Select a topic from the list to see the map containing that topic.

Use the **Clear** button to clear a find request.

**Documents button.** Pressing the **Documents** button displays a list of the documents containing the phrases represented by the selected nodes.

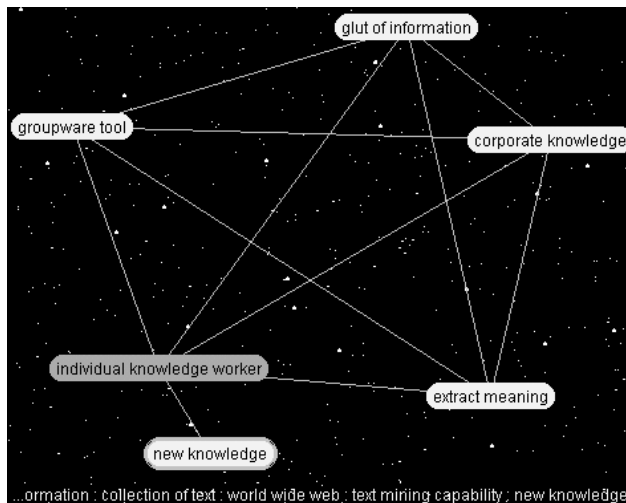
- If you select a single node, SemioMap displays the documents containing that topic.
- If you select multiple nodes (by pressing the Shift key when clicking on additional nodes), the document list represents the documents in which these topics occur together.
- If you select two unrelated nodes (nodes not linked with a line or in a cycle), you may see an empty document list, as no documents contain *both* of the selected topics.

## Understanding Map Shapes

The concept map is made up of many connected individual maps. (A map is the collection of nodes and relationships visible on the screen at one time.)

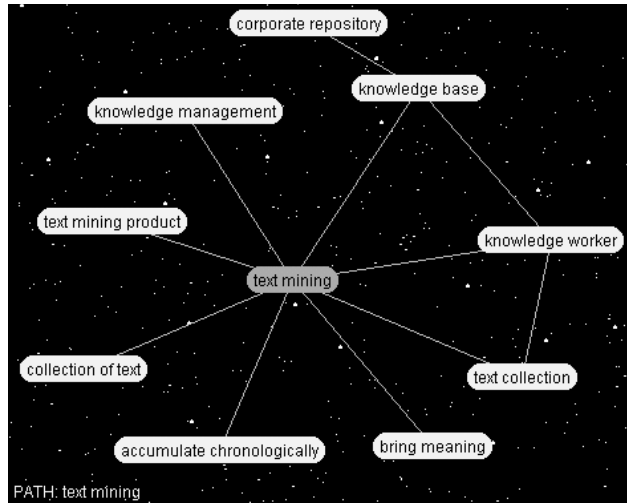
You can learn about the document collection from the shape of the maps.

**Cycles.** A cycle is a closely related groups of nodes, typically interconnected. A triangle of nodes is a small cycle. Nodes in a tightly connected cycle probably come from a very small set of documents, or a single document.



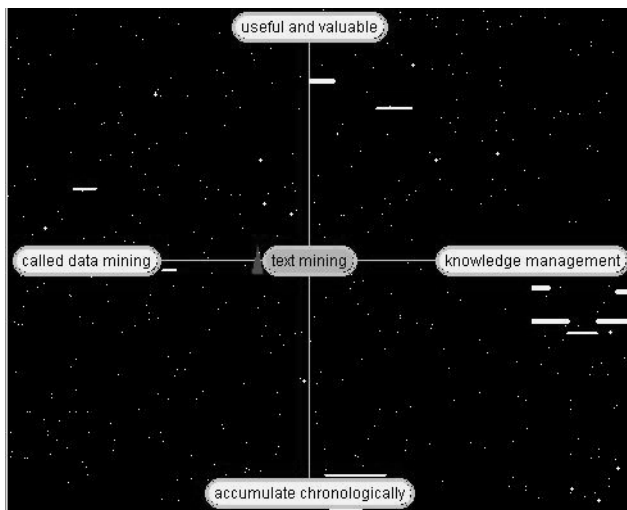
If you see highly connected cycles at a very general level in the concept map (level 1 or 2, for example), then this indicates a grouping of important topics in the document.

**Pivot nodes.** A pivot node is a central node that has single links to other, unconnected nodes, like the spokes of a wheel. A pivot node represents the common topic in a number of different discussions.



A pivot may bring together many separate ideas, often from diverse documents.

**Transitive links.** A *transitive links* connects groups of maps through external links. Transitive links display the “backbone” of the document collection.



---

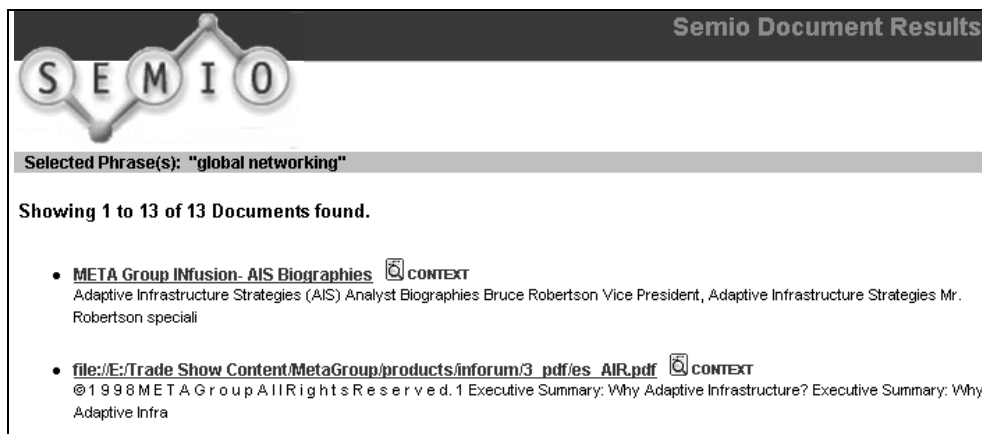
*“Text mining” is a transitive node, connecting several different topic areas.*

Chains of nodes showing transitive relationships show a “flow of thought” from one idea to the next.

## Viewing Source Documents

When you press the **Documents** button, the document list opens in a separate browser window. When you are done looking at source documents, you can either close this window or leave it open in the background. If you keep this window open in the background, it will be refreshed with a new document list each time you press the **Documents** button in the Discovery screen.

The document list looks like the following (although the page format may be customized for your installation):



Selecting the document's link displays the document in full.

- If the source is a web page, you travel to the web page.
- If the source is a file, the browser prompts you to launch the program in which the file was created.

If you select the **Context** link, you see a list of paragraphs from the document(s) that include the selected phrase (represented by the node you selected). The key phrases are highlighted.



---

# *Index*

## **C**

- catalog
  - SemioMap Discovery 28
- categorization 6
- concept maps 23, 24
  - navigating 30
- concepts 17
- Context link 19, 37
- cross references 18
- cycles 34

## **D**

- data sources 10
- databases 10, 28
- detail levels 25, 31
- details slide bar 31
- directory 5
- discovery 3, 4
- documents 19
  - context 37
  - Context link 19
  - viewing 37
- Documents button 33, 37

## **F**

- Find box 30, 32

## **H**

- hiding 19

## **I**

- index 5

## **K**

- key concepts 10, 12
- keyword search 3

## **L**

- links
  - transitive 36
- Lotus Notes 10

## **M**

- maps 25, 28, 29, 32
  - selecting 32
- Maps scroll box 30, 32

## **N**

- nodes 24
  - cycles 34
  - pivot 35
  - red arrows 31
  - satellite 31
  - shared 25, 31

## **O**

- ontology 10, 13
- Other Paths button 17

## **P**

- panning 25
- pivot nodes 35

## **R**

- Refine Search button 18, 20
- related terms 12
- relevance bar 19

## **S**

- search
  - keyword 2, 3
  - taxonomy 21
- Semio Tagger 10
- Semio Taxonomy 5
- SemioMap 5, 23
- SemioMap Discovery
  - catalog 28
  - screen 30
  - starting 27
- shared nodes 31
- Solaris
  - starting SemioMap Discovery 27
- source documents 37
- sources 10
  - displaying 37
- structured data 1
- subcategories 17
- Submit button 32

## **T**

- taxonomy 5, 10
  - concepts 17
  - multiple 11
  - navigating 15
- Other Paths button 17
- searching 21
- subcategories 17

---

- updating 11
- users 14
- transitive links 36

## **W**

- web sites 10
- Windows
  - starting SemioMap Discovery 27
- worlds 26, 32
  - selecting 32
- Worlds scroll box 30, 32

## **X**

- XML 10